

UNIVERSIDADE FEDERAL DO PARANÁ

ANTONIO CARLOS DA SILVA JÚNIOR

PROBABILIDADE E A DISTRIBUIÇÃO NORMAL MULTIVARIADA

CURITIBA

2022

ANTONIO CARLOS DA SILVA JÚNIOR

PROBABILIDADE E A DISTRIBUIÇÃO NORMAL MULTIVARIADA

Trabalho apresentado para compor a avaliação da disciplina de Probabilidade e Estatística Matemática I, ministrada pelo Prof. Dr. Anderson Ara no Programa de Pós Graduação em Métodos Numéricos em Engenharia, do Setor de Ciências Exatas da Universidade Federal do Paraná.

CURITIBA

2022

SUMÁRIO

1	INTRODUÇÃO À PROBABILIDADE	3
1.1	VARIÁVEIS ALEATÓRIAS	5
2	DISTRIBUIÇÃO NORMAL MULTIVARIADA	7
2.1	DISTRIBUIÇÃO NORMAL BIVARIADA	8
2.2	APLICAÇÕES	10
	REFERÊNCIAS	14

1 INTRODUÇÃO À PROBABILIDADE

Probabilidade é o estudo da aleatoriedade e da incerteza. Para chegarmos a uma definição mais sólida, é necessário definirmos antes conceitos de espaço amostral e eventos de um experimento.

Espaço amostral é o conjunto de todos os resultados possíveis de um experimento aleatório, e será representado pela letra Ω neste trabalho. Por exemplo, se um experimento consiste no lançamento de um dado, então o espaço amostral é formado por 6 elementos, que correspondem às faces do dado. Ou seja, $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Evento é um subconjunto do espaço amostral composto pelos resultados de interesse em um experimento. Por exemplo, no experimento do lançamento de um dado, se estamos interessados na ocorrência de números pares, podemos dizer que o evento A é composto pelos elementos $\{2, 4, 6\}$. E então observamos que $A \subset \Omega$.

Dois ou mais eventos podem formar uniões, intersecções e complementos. Ainda com base no exemplo do lançamento de dados, adicionaremos o evento B , composto pelos elementos menores ou iguais a 3. Ou seja, $B = \{1, 2, 3\}$.

- A **união** entre A e B resulta em um novo evento contendo todos os elementos de A , B ou de ambos. Por exemplo, $A \cup B = \{1, 2, 3, 4, 6\}$;
- A **intersecção** entre A e B resulta em um novo evento contendo todos os elementos que estejam simultaneamente nos dois eventos. Por exemplo, $A \cap B = \{2\}$;
- O **complemento** de A , é um evento composto pelos elementos que não estão presentes em A . Por exemplo, $A^c = \{1, 3, 5\}$.

Mais dois tipos de eventos são importantes para definirmos probabilidade: os independentes e os mutuamente exclusivos. A e B são independentes quando a probabilidade de ocorrência de B não for condicional à probabilidade de ocorrência de A . Por exemplo, em um experimento composto pelo lançamento de duas moedas, a probabilidade de sair cara da segunda moeda não depende do resultado do primeiro lançamento. A e B são mutuamente exclusivos quando não podem ocorrer simultaneamente. Por exemplo, considerando o lançamento de um dado, o resultado não pode ser um número par e um número ímpar ao mesmo tempo.

Na **definição clássica**, probabilidade é a razão do número de resultados favoráveis a um evento (m) e o número de casos possíveis (n). Por exemplo, em uma rifa com 100 números (n), a probabilidade de uma pessoa que assinou 2 números (m) ser sorteada é $2/100 = 0.02$.

Se fizermos uso da definição clássica para calcular, por exemplo, a probabilidade de sucesso de um foguete tripulado lançado à lua, sendo sucesso o foguete chegar à lua com os tripulantes vivos, teremos o seguinte: $\Omega = \{1, 0\}$, $A = \{1\}$, $P(A) = m/n = 1/2 = 0.5$, o que soa como um absurdo.

A definição clássica é aplicada quando todos os casos possíveis são igualmente prováveis, o que não é verdade no exemplo do foguete. Podemos tentar por meio da **definição frequentista** de probabilidade. Se o lançamento do foguete for repetido n vezes exatamente sob as mesmas condições, e m for o número de sucessos entre as n repetições. Então, se n for suficientemente grande, m/n é uma boa aproximação da probabilidade de sucesso.

Não parece uma boa ideia fazer uso da definição frequentista para calcular a probabilidade de sucesso no exemplo do foguete. No entanto, também existe a **definição axiomática** de probabilidade. Pela definição axiomática, a probabilidade do evento A em um espaço amostral Ω , denotada por $P(A)$, é um número real no qual os seguintes axiomas são obedecidos:

1. $0 \leq P(A) \leq 1$;
2. $P(\Omega) = 1$;
3. Se A e B são eventos mutuamente exclusivos, $P(A \cup B) = P(A) + P(B)$.

A definição axiomática não nos diz o que é probabilidade, nem como calculá-la, mas permite, por meio dos três axiomas apresentados, desenvolver uma série de propriedades para podermos calcular probabilidades de eventos complexos com base na probabilidade de eventos elementares. Portanto, retornando ao exemplo do foguete, mesmo sem nunca ter realizado tal experimento, é possível calcular a probabilidade de sucesso por meio das propriedades desenvolvidas a partir da definição axiomática.

Ainda existe a **definição subjetiva** de probabilidade. Esta definição é adotada quando não temos outra forma de calcular a probabilidade. Por exemplo, dois eleitores podem divergir quanto a probabilidade de um determinado governante ser bem-sucedido durante o governo. Esta divergência ocorre em função das estimativas, conceitos prévios e informações que cada um dos eleitores possuem.

Sendo assim, probabilidade pode ser definida como a medida de informação ou crença sobre a ocorrência de um evento.

Nota: Esta sessão foi estruturada com base nos livros de Ross (2010), Pinheiro et al. (2012) e Fávero e Belfiore (2017), e nas aulas do professor José Luiz Padilha (UFPR).

1.1 VARIÁVEIS ALEATÓRIAS

Ross (2010) explica o que são variáveis aleatórias por meio do seguinte exemplo: dado um experimento que consiste no lançamento de 3 moedas, com 0 representando cara e 1 representando coroa. Se X representar o número de caras que aparecem no resultado, então X é uma variável aleatória que pode assumir os valores 0 quando o resultado for $\{0, 0, 0\}$, 1 quando o resultado for $\{(0, 0, 1), (0, 1, 0), (1, 0, 0)\}$, 2 quando $\{(0, 1, 1), (1, 0, 1), (1, 1, 0)\}$ ou 3 quando $\{(1, 1, 1)\}$. As probabilidades de cada um dos resultados de X ficam definidas como:

- $P(X = 0) = \frac{m}{n} = \frac{1}{8};$
- $P(X = 1) = \frac{m}{n} = \frac{3}{8};$
- $P(X = 2) = \frac{m}{n} = \frac{3}{8};$
- $P(X = 3) = \frac{m}{n} = \frac{1}{8}.$

Observe que $P(\Omega) = P\left[\bigcup_{i=0}^3 (X = i)\right] = \sum_{i=0}^3 P(X = i) = 1.$

De acordo com Ross (2010), geralmente, diante de um experimento, o pesquisador está interessado em alguma função do resultado e não especificamente no resultado em si. No exemplo anterior, o interesse está no número de caras e não nos resultados individuais que compõem a soma de caras. Estas funções reais definidas no espaço amostral são variáveis aleatórias.

As variáveis aleatórias podem ser discretas ou contínuas. As variáveis aleatórias discretas assumem valores contáveis, ao passo que as variáveis aleatórias contínuas assumem valores incontáveis. Por exemplo, o tempo de vida de um determinado equipamento pode ser representado por meio de uma variável aleatória contínua. Segundo Ross (2010), X é uma variável aleatória contínua se existir uma função densidade de probabilidade, ou seja, uma função não negativa $f(x)$ definida para todo real $x \in (-\infty, \infty)$ em que, para qualquer conjunto B de números reais,

$$P(X \in B) = \int_B f(x)dx. \quad (1.1)$$

Em outras palavras, a equação 1.1 indica que a probabilidade de que X esteja em B pode ser calculada por meio da integração da função densidade de probabilidade ao longo do conjunto B . Como X deve assumir algum valor, temos

$$P[X \in (-\infty, \infty)] = \int_{-\infty}^{\infty} f(x)dx = 1, \quad (1.2)$$

$$P(a \leq X \leq b) = \int_a^b f(x)dx, \quad (1.3)$$

$$P(X = a) = \int_a^a f(x)dx = 0, \quad (1.4)$$

$$P(X \leq a) = P(X < a) = \int_{-\infty}^a f(x)dx. \quad (1.5)$$

Se X é uma variável aleatória contínua com função densidade de probabilidade $f(x)$, então a esperança, ou valor esperado de X , é definido como

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx, \quad (1.6)$$

, variância de X definida como

$$\text{Var}(X) = E[(X - \mu)^2] = E(X^2) - [E(X)]^2, \quad (1.7)$$

em que $\mu = E(X)$, e função geradora de momentos definida como

$$M(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx}f(x)dx \quad (1.8)$$

para todos os valores reais de t .

O comportamento das variáveis aleatórias pode ser descrito por meio das distribuições de probabilidades. Entre as variáveis aleatórias contínuas, as distribuições abaixo estão entre as mais conhecidas na literatura:

- Distribuição Uniforme Contínua,
- Distribuição Exponencial,
- Distribuição Gama,
- Distribuição Normal.

2 DISTRIBUIÇÃO NORMAL MULTIVARIADA

Seja X uma variável aleatória com função densidade de probabilidade $f(x)$, função de distribuição acumulada $F(x) = P(X \leq x)$, função geradora de momentos $M_X(t)$, esperança $E(X)$ e variância $\text{Var}(X)$, X tem distribuição normal univariada, ou seja, $X \sim N(\mu, \sigma^2)$, se sua função densidade de probabilidade é dada por

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right\}, \quad (2.1)$$

com $-\infty < x < \infty$, $\mu = E(X)$ e $\sigma^2 = \text{Var}(X)$.

A distribuição normal univariada pode ser generalizada para a distribuição normal multivariada, para o caso no qual se trabalha com duas ou mais variáveis aleatórias simultaneamente. Para tornar mais clara a generalização que será apresentada a seguir, observe que a equação 2.1 pode ser rearranjada como

$$f(x) = \frac{1}{(2\pi)^{\frac{1}{2}}(\sigma^2)^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2}(x - \mu)(\sigma^2)^{-1}(x - \mu) \right\}. \quad (2.2)$$

Seja $\mathbf{X} = (X_1, \dots, X_p)^T$ uma coleção de variáveis aleatórias, denominado vetor aleatório, com função densidade de probabilidade $f(\mathbf{x})$, função de distribuição acumulada $F(\mathbf{x}) = P(X_1 \leq x_1, \dots, X_p \leq x_p)$, função geradora de momentos $M_{\mathbf{X}}(\mathbf{t})$, vetor de esperanças $E(\mathbf{X})$ e matriz de variâncias e covariâncias $\text{Cov}(\mathbf{X})$, \mathbf{X} tem distribuição normal multivariada, ou seja, $\mathbf{X} \sim N_p(\mu, \Sigma)$, se sua função densidade de probabilidade é dada por

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{p}{2}} |\Sigma|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2}(\mathbf{x} - \mu)^T \Sigma^{-1}(\mathbf{x} - \mu) \right\}, \quad (2.3)$$

com $-\infty < x_i < \infty$, $i = 1, 2, \dots, p$, $\mu_{p \times 1} = E(\mathbf{X})$ e $\Sigma_{p \times p} = \text{Cov}(\mathbf{X})$. Assumindo que a matriz simétrica Σ é positiva definida, a expressão 2.4 é a Distância de Mahalanobis entre \mathbf{x} e μ .

$$(\mathbf{x} - \mu)^T \Sigma^{-1}(\mathbf{x} - \mu). \quad (2.4)$$

De acordo com Johnson e Wichern (2014), as propriedades abaixo são válidas para $\mathbf{X} \sim N_p(\mu, \Sigma)$:

1. Qualquer combinação linear das componentes de \mathbf{X} são normalmente distribuídas;

2. Todos os subconjuntos das componentes de \mathbf{X} têm distribuição normal (univariada ou multivariada);
3. Zero covariância implica que as respectivas componentes de \mathbf{X} são independentes;
4. A distribuição condicional das componentes de \mathbf{X} são normal (univariada ou multivariada).

2.1 DISTRIBUIÇÃO NORMAL BIVARIADA

Na equação 2.3, quando $\mathbf{X} = (X_1, X_2)^T$ e, conseqüentemente, $p = 2$, temos a distribuição normal com duas dimensões ou distribuição normal bivariada. Neste caso é possível reescrever a matriz de variâncias e covariâncias em termos de correlação, como demonstrado em 2.5.

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix} = \begin{bmatrix} \rho_{11} & \rho_{12}\sqrt{\sigma_{11}\sigma_{22}} \\ \rho_{12}\sqrt{\sigma_{11}\sigma_{22}} & \rho_{22} \end{bmatrix}. \quad (2.5)$$

Em conseqüência, o determinante e a inversa de Σ ficam definidos por 2.6 e 2.7, respectivamente.

$$|\Sigma| = \sigma_{11}\sigma_{22} - \sigma_{12}^2 = \rho_{11}\rho_{22}(1 - \rho_{12}^2) \quad (2.6)$$

$$\Sigma^{-1} = \frac{1}{1 - \rho_{12}^2} \begin{bmatrix} \frac{1}{\rho_{11}} & -\frac{\rho_{12}}{\sqrt{\rho_{11}\rho_{22}}} \\ -\frac{\rho_{12}}{\sqrt{\rho_{11}\rho_{22}}} & \frac{1}{\rho_{22}} \end{bmatrix} \quad (2.7)$$

Observe em 2.6 que se a correlação entre as variáveis aleatórias X_1 e X_2 for perfeita, ou seja, $\rho_{12} = 1$, o determinante da matriz de variâncias e covariâncias será nulo. Observe em 2.3 que $|\Sigma| = 0$ anula a função, resultando na não existência de uma distribuição normal multivariada.

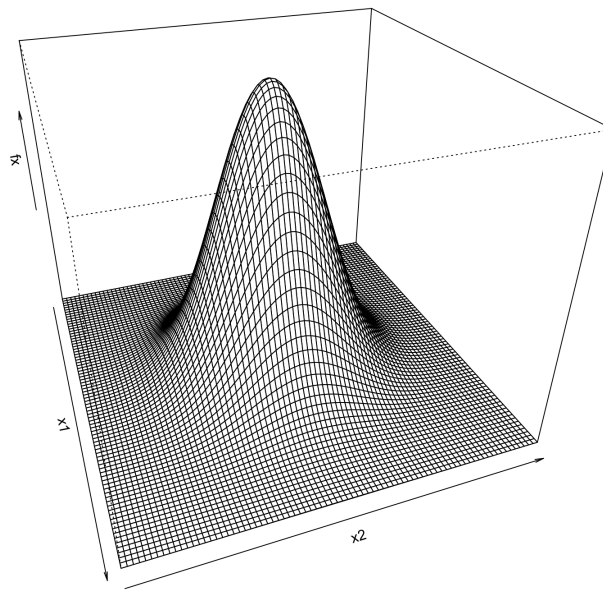
Por fim, a diferença $(\mathbf{x} - \mu)$ é obtida por meio de uma simples operação de subtração entre matrizes demonstrada em 2.8, e então, substituindo 2.5, 2.6, 2.7 e 2.8 em 2.3, obtemos 2.9.

$$(\mathbf{x} - \mu) = \begin{bmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{bmatrix}. \quad (2.8)$$

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{p}{2}} \sqrt{\sigma_{11}\sigma_{22}(1-\rho_{12}^2)}} \exp \left\{ -\frac{1}{2(1-\rho_{12}^2)} \left[\left(\frac{x_1 - \mu_1}{\sqrt{\sigma_{11}}} \right)^2 + \left(\frac{x_2 - \mu_2}{\sqrt{\sigma_{22}}} \right)^2 - 2\rho_{12} \left(\frac{x_1 - \mu_1}{\sqrt{\sigma_{11}}} \right) \left(\frac{x_2 - \mu_2}{\sqrt{\sigma_{22}}} \right) \right] \right\}. \quad (2.9)$$

Com a expressão 2.9 podemos discutir algumas propriedades da distribuição normal multivariada. Comentamos anteriormente que se a correlação entre as variáveis aleatórias X_1 e X_2 for perfeita, não há distribuição normal univariada. Agora observamos que se não houver correlação entre X_1 e X_2 , ou seja, $\rho_{12} = 0$, que $f(\mathbf{x})$ pode ser escrita como o produto entre as densidades univariadas de X_1 e X_2 (2.1) e X_1 e X_2 são independentes. As figuras 1 e 2 exibem a densidade normal bivariada com $\rho_{12} = 0$ e $\rho_{12} = 0.9$, respectivamente.

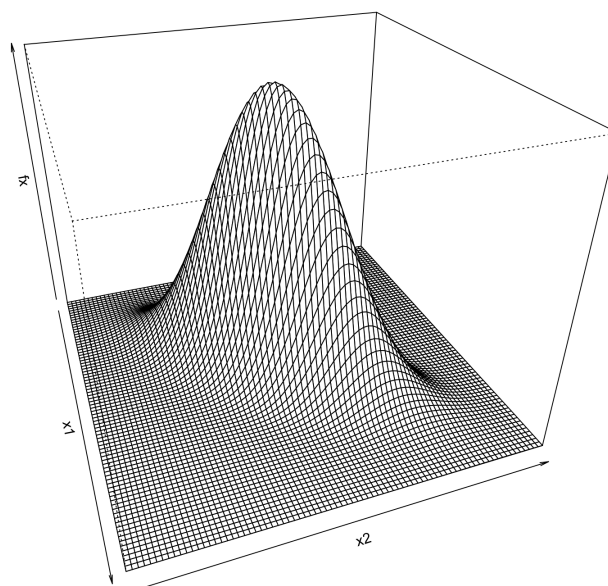
FIGURA 1 – DENSIDADE NORMAL BIVARIADA COM $\sigma_{11} = \sigma_{22}$ E $\rho_{12} = 0$



FONTE: O autor

A densidade da normal multivariada é constante em superfícies em que

$$(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) = c^2. \quad (2.10)$$

FIGURA 2 – DENSIDADE NORMAL BIVARIADA COM $\sigma_{11} = \sigma_{22}$ E $\rho_{12} = 0.9$ 

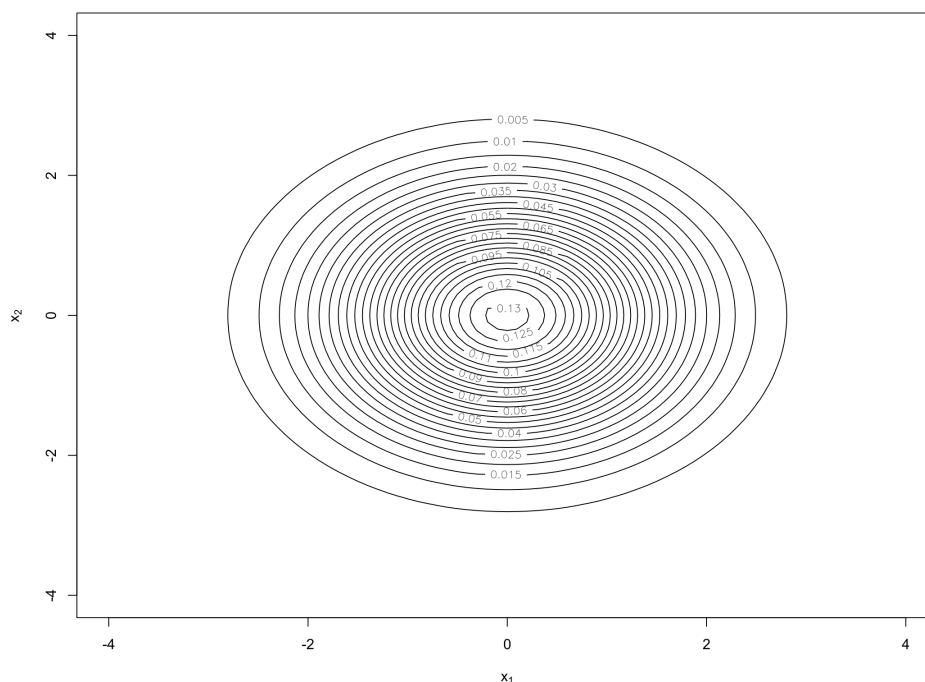
FONTE: O autor

e o conjunto dos x que satisfazem a equação 2.10 são os que compõem o **contorno de densidade de probabilidade constante**. Os contornos de densidade constante da normal multivariada são elipsoides com centro em μ e eixos em $\pm c\sqrt{\lambda_i}e_i$ em que (λ_i, e_i) é um par de autovalor-autovetor de Σ^{-1} .

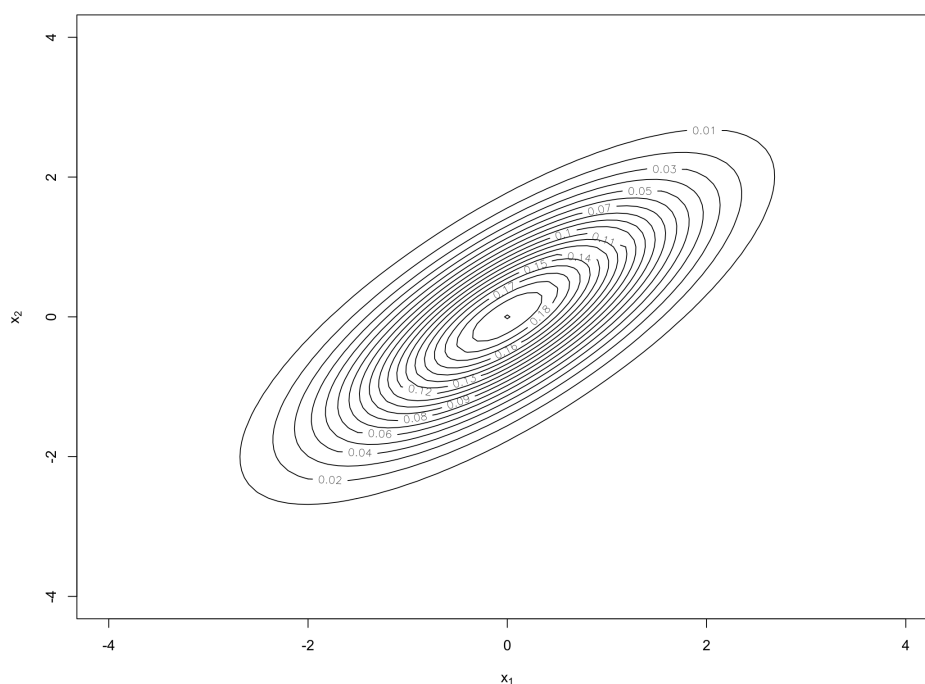
As figuras 3 e 4 exibem os contornos para os casos em que os elementos são não correlacionados ($\rho_{12} = 0$) e correlacionados ($\rho_{12} \neq 0$), respectivamente. Observe que para o caso não correlacionado, os autovetores de Σ^{-1} são paralelos aos eixos x_1 e x_2 , ao passo que para o caso correlacionado, os autovetores apresentam uma inclinação.

2.2 APLICAÇÕES

De acordo com Furtado (1996), a maioria das técnicas multivariadas parte do pressuposto de que os dados seguem uma distribuição normal multivariada, o que faz com que a distribuição normal multivariada seja de suma importância na análise multivariada. Na prática raramente os dados terão distribuição normal multivariada perfeita, mas uma aproximação em muitos casos é suficiente para facilitar todo o

FIGURA 3 – CONTORNOS ELÍPTICOS DA NORMAL BIVARIADA COM $\sigma_{11} = \sigma_{22}$ E $\rho_{12} = 0$ 

FONTE: O autor

FIGURA 4 – CONTORNOS ELÍPTICOS DA NORMAL BIVARIADA COM $\sigma_{11} = \sigma_{22}$ E $\rho_{12} = 0$ 

FONTE: O autor

tratamento matemático. Furtado (1996) menciona que devido ao teorema do limite central, a distribuição normal multivariada se aproxima da distribuição amostral de

muitas estatísticas multivariadas, independentemente da distribuição original.

Uma conhecida técnica multivariada é a **Análise de Discriminante**. Segundo Tabachnick e Fidell (2013), o objetivo desta técnica é prever a associação do grupo a partir de um conjunto de preditores. Por exemplo, é possível um diagnóstico diferenciado entre um grupo de crianças sem deficiência, um grupo de crianças com dificuldade de aprendizado e um grupo com transtorno com transtorno emocional de forma confiável a partir de um conjunto de pontuações de testes psicológicos? Tabachnick e Fidell (2013) mencionam ainda a necessidade de que as pontuações nos preditores sejam amostras independentes e aleatórias de uma população e que a distribuição amostral de qualquer combinação linear de preditores seja normalmente distribuídas. Em outras palavras, as autoras se referem à suposição de normalidade multivariada. AbuZeina e Al-Anzi (2018) utilizaram a Análise de Discriminante Linear para classificação de textos em Árabe e tiveram resultados promissores para aplicações no campo da mineração de texto.

Outra técnica multivariada é a **Análise de Variância Multivariada**, também conhecida como MANOVA, que é uma generalização da ANOVA para casos várias variáveis dependentes. Tabachnick e Fidell (2013) exemplificam da seguinte forma: imagine que um pesquisador esteja interessado no efeito de diferentes tipos de tratamento para diversos tipos de ansiedade, sendo que a variável independente é composta por três níveis de tratamento. Após a atribuição aleatória dos indivíduos a tratamentos e um período subsequente de tratamento, os indivíduos são medidos quanto aos três tipos de ansiedade. As pontuações em todas as três medidas servem como variáveis dependentes. O objetivo desta técnica é avaliar se uma combinação das três medidas varia em função do tratamento. A diferença da MANOVA para Análise de Discriminante é que enquanto esta enfatiza as diferenças médias e significância estatística, a Análise de Discriminante enfatiza a previsão de associação ao grupo e as dimensões nas quais os grupos diferem. De acordo com Tabachnick e Fidell (2013), os testes de significância para MANOVA são baseados na distribuição normal multivariada, uma vez que a normalidade multivariada implica que as as distribuições amostrais das médias das variáveis dependentes em todas as combinações lineares sejam normalmente distribuídas. Com F univariado e amostras suficientemente grandes, o teorema do limite central sugere que as distribuições amostrais das médias se aproxima da normalidade mesmo quando os escores originais não. Johnson, Lloyd et al. (2007) demonstraram o poder e a versatilidade da Análise de Variância Multivariada por meio de duas técnicas metabolômicas diferentes. Os autores encontraram efeitos principais e interações significativas no experimentos, permitindo uma interpretação mais completa e abrangente do que com a Análise de Componentes Principais.

De acordo com Ullman (2013), a **Modelagem de Equações Estruturais** é

um conjunto de métodos estatísticos que permitem avaliar o relacionamento entre uma ou mais variáveis independentes e uma ou mais variáveis dependentes, podendo as variáveis ser contínuas ou discretas, fatores ou métricas. A maioria dos métodos usados na Modelagem de Equações Estruturais assumem normalidade multivariada. Por exemplo, método de estimação da **Máxima Verossimilhança**, que faz uso de um procedimento iterativo para minimização da função de verossimilhança, e, segundo Anderson e Gerbing (1982), sob a suposição de normalidade multivariada este método se apresenta de forma consistente, eficiente e sem vieses de estimativas. O método dos **Mínimos Quadrados Generalizados**, conhecido no campo da Econometria, é aplicado na estimação de parâmetros desconhecidos em um modelos de regressão linear, em caso de heterocedasticidade ou quando há correlação entre os resíduos. Segundo Amorim et al. (2012), sob a suposição de normalidade multivariada, o método dos Mínimos Quadrados Generalizados é assintoticamente normal e eficiente. Liu e Liu (2022) apresentaram um novo método para estimação de parâmetros em equações diferenciais baseado no método da Máxima Verossimilhança. Bai, Choi e Liao (2020) utilizaram a estimação por Mínimos Quadrados Generalizados para modelos de dados de painel linear e concluíram que o estimador proposto é mais eficiente quando os Mínimos Quadrados Ordinários, quando há heterocedasticidade, correlações seriais e transversais.

REFERÊNCIAS

ABUZEINA, Dia; AL-ANZI, Fawaz S. Employing fisher discriminant analysis for Arabic text classification. **Computers and Electrical Engineering**, Elsevier BV, v. 66, p. 474–486, fev. 2018. DOI: 10.1016/j.compeleceng.2017.11.002. Citado 1 vez na página 12.

AMORIM, Leila D. A. F. et al. **Modelagem com Equações Estruturais: Princípios Básicos e Aplicações**. [S.l.: s.n.], 2012. Disponível em: <https://repositorio.ufba.br/bitstream/ri/17684/1/ebook_SEM_2012.pdf>. Acesso em: 10 jul. 2022. Citado 1 vez na página 13.

ANDERSON, James C.; GERBING, David W. Some Methods for Respecifying Measurement Models to Obtain Unidimensional Construct Measurement. **Journal of Marketing Research**, SAGE Publications, v. 19, n. 4, p. 453–460, nov. 1982. DOI: 10.1177/002224378201900407. Citado 1 vez na página 13.

BAI, Jushan; CHOI, Sung Hoon; LIAO, Yuan. Feasible generalized least squares for panel data with cross-sectional and serial correlations. **Empirical Economics**, Springer Science e Business Media LLC, v. 60, n. 1, p. 309–326, nov. 2020. DOI: 10.1007/s00181-020-01977-2. Citado 1 vez na página 13.

FÁVERO, Luiz Paulo; BELFIORE, Patrícia. **Manual de análise de dados: Estatística e modelagem multivariada com Excel®, SPSS® e Stata®**. [S.l.]: Elsevier Brasil, 2017. Citado 1 vez na página 4.

FURTADO, Daniel. **Análise Multivariada**. [S.l.: s.n.], 1996. Disponível em: <<http://ftp.demec.ufpr.br/disciplinas/TM788/Daniel%20Furtado%20Ferreira/>>. Acesso em: 9 jul. 2022. Citado 2 vezes nas páginas 10, 11.

JOHNSON, Helen E.; LLOYD, Amanda J. et al. The application of MANOVA to analyse Arabidopsis thaliana metabolomic data from factorially designed experiments. **Metabolomics**, Springer Science e Business Media LLC, v. 3, n. 4, p. 517–530, jun. 2007. DOI: 10.1007/s11306-007-0065-3. Citado 1 vez na página 12.

JOHNSON, Richard; WICHERN, Dean. **Applied Multivariate Statistical Analysis**. [S.l.]: Pearson, 2014. Citado 1 vez na página 7.

LIU, Yang; LIU, Baoding. Estimating unknown parameters in uncertain differential equation by maximum likelihood estimation. **Soft Computing**, Springer Science e Business Media LLC, v. 26, n. 6, p. 2773–2780, jan. 2022. DOI: 10.1007/s00500-022-06766-w. Citado 1 vez na página 13.

PINHEIRO, João Ismael D. et al. **Probabilidade e Estatística: Quantificando a incerteza**. [S.l.]: Elsevier Brasil, 2012. Citado 1 vez na página 4.

ROSS, Sheldon. **Probabilidade: Um curso moderno com aplicações**. [S.l.]: Bookman, 2010. Citado 4 vezes nas páginas 4, 5.

TABACHNICK, Barbara G.; FIDELL, Linda S. **Using Multivariate Statistics**. [S.l.]: Pearson, 2013. Citado 4 vez na página 12.

ULLMAN, Jodie B. **Using Multivariate Statistics**. [S.l.]: Pearson, 2013. Citado 1 vez na página 12.